Getting to reproducible antibodies: the rationale for sequenced recombinant characterized reagents

Bradbury, Andrew R M ; Plückthun, Andreas

DOI: https://doi.org/10.1093/protein/gzv051

Protein Engineering, Design & Selection, 2015, vol. 28 no. 10, pp. 303–305
doi: 10.1093/protein/gzv051
Commentary

Commentary

Getting to reproducible antibodies: the rationale for sequenced recombinant characterized reagents

Andrew R.M. Bradbury[1,*] and Andreas Plückthun[2,*]

[1]Los Alamos National Laboratory, Los Alamos, NM 87545, USA, and [2]Department of Biochemistry, University of Zurich, Zurich CH-8057, Switzerland

*To whom correspondence should be addressed. E-mail: amb@lanl.gov (A.R.M.B); plueckthun@bioc.uzh.ch (A.P)

Would you work with a gene, vector or oligonucleotide of unknown sequence? Today, the majority of scientists would answer with a resounding 'No'. Biology is becoming more quantitative, digital and defined. The availability of genomic sequences in public databases allows the same genes to be effectively and reproducibly synthesized and studied in different laboratories everywhere: where a sequence is available there is no uncertainty. The genes, vectors, RNAs or oligonucleotides I order will be indistinguishable from the ones you describe in your experiments. This allows me to attempt to reproduce your experiments, and while there may be differences in the protocols we use, at least we can be sure that the underlying biological materials are essentially identical. However, the situation is very different when it comes to the specific detection of molecules in biological samples using 'specific' probes.

The most widely used specific binding reagents are synthetic nucleic acid probes based on sequence knowledge. These are able to recognize RNA or DNA by hybridization with enormous specificity and affinity, and rely on well-understood base complementarity for their high specificity. For that reason, they are reproducible between laboratories, and even partial cross-reactivities are predictable to some degree. Nowadays, the sequences of oligonucleotides used in an experimental study are typically listed in a Supplementary Table in each publication, allowing relatively straightforward experimental duplication.

For historical, as well as practical, reasons, antibodies are by far the most widely used class of specific detection reagents for essentially all other target classes, particularly proteins. Polyclonal antisera have been used in research for nearly a century, and monoclonal antibodies for four decades (Köhler and Milstein, 1975). However, although these reagents have been instrumental in addressing numerous biomedical research questions, they are never defined at the molecular level (we do not consider here the use of antibodies as biological pharmaceuticals, as these are all highly quality-controlled, recombinant reagents that have been exquisitely characterized). Animal-derived reagent antibodies are the main subject of this editorial.

Without access to alternatives, researchers have become accustomed to—and usually do not question—the inadequate definition and characterization of these traditional reagents, even in an era when working with oligonucleotides, genes, vectors or even *genomes* of unknown sequence is inconceivable. Modern biomedical and clinical research relies on specific, high-affinity detection reagents that are functional in complex environments. They provide information on whether a particular component is present in a biological sample, how much of it there is, where it is found, and with which other macromolecules it interacts. The nature of the specific detection reagent in molecular terms (i.e. whether it is an antibody, another scaffold or an aptamer) is less important than its quality, assessed in terms of specificity, epitope recognized, affinity and functionality in different assays, and the ability to describe it sufficiently well that other scientists can reproducibly use the same reagent.

Progress in research relies on reproducibility—the generation of reliable results upon which future studies can be dependably based. However, many experts, including NIH Director, Francis Collins, believe our 'system for ensuring reproducibility of biomedical research is failing' (Collins and Tabak, 2014). Clearly, the ability to repeat experiments with reagents identical to those used in previous publications is an essential part of creating a successful and reproducible biomedical research environment.

For all those reagents defined at the molecular or sequence level described above, this is an achievable goal, and can be eliminated as a source of irreproducibility. However, when it comes to antibodies, the situation is very different for a number of different reasons. Obviously, the molecular definition of polyclonal antibodies will remain impossible at a practical level, and while the genes of monoclonal antibodies can be sequenced and therefore completely defined, this is rarely carried out.

In addition to the problem of inadequate definition, the situation is further clouded by the fact that it is projected no more than 35–50% of commercial antibodies actually recognize their targets with the claimed specificity (Berglund *et al.*, 2008; Slaastad *et al.*, 2011); although this projection is unevenly distributed, with some manufacturers producing consistently high, and others consistently low, quality antibodies (Berglund *et al.*, 2008; Bordeaux *et al.*, 2010). This is further exacerbated by intrinsic lot-to-lot variability, particularly

with polyclonal antibodies, and two widely unrecognized market practices: data sheets often do not even correspond to the lot sold (Voskuil, 2014), and the same original antibodies are often sold by different providers under different labels (Alm *et al.*, 2014; Voskuil, 2014). The problem is probably best summarized thus: antibodies sold as different are often identical, while antibodies sold as identical are often different (thanks to Natalie de Souza (editor Nature Methods) for this pithy insightful observation), and the customer does not know which is which.

Although most researchers consider monoclonal antibodies to be perfectly defined reagents with single specificities and so free of these problems, hybridomas frequently secrete more than one light and/or heavy chain (Ruberti *et al.*, 1994; Zack *et al.*, 1995; Blatt *et al.*, 1998). This makes it difficult to assess whether the binding properties of a monoclonal antibody preparation are intrinsic to the monoclonal antibody molecule itself, or the result of additional specificities caused by the presence of additional chains. This can only be determined when all expressed V genes are cloned and assessed for their binding properties in all possible combinations. Furthermore, hybridomas are known to lose expression, or mutate, requiring frequent recloning and testing. Given these caveats, the identity, reproducibility (and monoclonality) of two purportedly identical monoclonal antibody samples derived from hybridomas cannot be clearly established.

Part of the problem is that protein–protein interactions are not nearly as predictable as base-pairing interactions in nucleic acids, making the experimental confirmation of antibody specificity prior to use essential (Bordeaux *et al.*, 2010; Bourbeillon *et al.*, 2010). However, even when antibodies do recognize their targets specifically, there are additional complexities. Antibodies are described in terms of the targets they (should) recognize, rather than in terms of their identity: definition of a reagent by what it *does*, rather than what it *is*, continues to be one of the major problems in the antibody field. Unfortunately, the first definition is not usually given with enough accuracy, and the second not at all. Antibodies specifically recognizing the same target protein may recognize different epitopes (linear or conformational), or different isoforms, of that the same target. The epitopes recognized by antibodies raised against peptides may be known, but are often not disclosed for commercial reasons. Thus, users normally have no way of knowing whether antibodies and/or their recognition specificities are identical, and this can lead to different results in matching experimental set-ups, even for antibodies that are truly specific for their targets.

This problem can only be overcome if antibodies, once characterized, become uniquely identifiable and their identities verifiable. As sequence is the defining characteristic of modern biology, we propose that the time has come to apply the gene-based paradigm to antibodies. The sequence is, after all, a unique bar code.

In the short term, in order to raise the quality of antibodies used in biomedical research (to reiterate, we are excluding therapeutic antibodies, which are excellent), it will be indispensable to develop widely adopted standards and best practices for antibody characterization and validation. However, this is necessary but not sufficient: in the long term, biological research should plan for the introduction of sequenced recombinant antibodies (or other binding molecules) as research reagents. Sequences should be made publicly available so that researchers can all use identical reproducible reagents (Bradbury and Plückthun, 2015a,b). Such reagents will, of course, require the same diligent characterization and validation as traditional antibodies—knowledge of the sequence is clearly independent of and additional to this requirement. However, unlike animal-derived antibodies, where this should be carried out prior to each use, extensive characterization of recombinant sequenced antibodies will be required only once (Bradbury and Plückthun, 2015a,b). Once binding and the concentration of active molecules is confirmed (using appropriate positive controls), it can be safely assumed that antibody binding properties will be essentially identical, providing they are produced under uniform conditions (it should be noted that recombinant antibodies expressed in different hosts can have different post-translational modifications. This can affect the homogeneity of populations of recombinant antibodies of the same sequence and contribute to differences in activity, hence the need for uniformity in production and standard operating procedures).

Even though antibodies are widely considered to be an important contributor to the irreproducibility problem plaguing biomedical research (Begley and Ellis, 2012; Begley, 2013; Bradbury and Plückthun, 2015a,b), it may be challenging to convince scientists to switch from the animal-derived antibodies they are familiar with to recombinant sequenced binders, even if this does improve reproducibility between different scientific studies. This indicates the need for collaboration between funding agencies and journals. Just as it is now impossible to publish the structure of a protein without depositing the coordinates in the PDB, so it should become impossible in the future to describe the use of an antibody that has not been defined at the sequence level—simply so that others can repeat the exact same experiment.

Commercial antibody providers are reluctant to release the sequences of the monoclonal antibodies they derive (Polakiewicz, 2015), regarding them (not unreasonably) as valuable proprietary information. This indicates the likely need for funding from public (or public–private partnerships) to derive and characterize antibodies initially. In the short term, we anticipate that valuable monoclonal antibody specificities should be cloned, sequenced and produced recombinantly, so avoiding the problems described above. In the longer term, we would expect recombinant reagents to be produced directly using display technologies. While antibodies have been widely used for decades, there are many new classes of affinity reagents on the horizon. These include recombinant antibodies created using natural or synthetic sources (Bradbury *et al.*, 2011), other recombinant protein scaffolds (e.g. DARPINs, anticalins, affibodies, fibronectin domains) (Binz *et al.*, 2005; Skerra, 2007, Veesler *et al.*, 2009), in which binding is mediated by diversity introduced on one surface that interacts with targets of interest; and nucleic acid-based affinity reagents (aptamers and SOMAmers), in which globular binding structures are provided by intricate three-dimensional folding motifs unlike natural nucleic acid folds (Davies *et al.*, 2012).

Using sequence to define these different recombinant reagents will allow them to compete in the research market place. Automation and molecular engineering, including streamlining of characterization, is just beginning, and expected to lead to significant cost reductions. Companies would then presumably compete on the quality, quantity or price of the binding reagents they produce, based on their publicly known sequences. The virtuous circle created by the need to disclose specific binder sequences will ensure that the poorly characterized irreproducible antibody will become a thing of the past.

References

Alm,T., von Feilitzen,K., Lundberg,E., Sivertsson,A. and Uhlen,M. (2014) *J. Proteome Res.*, **13**, 1669–1676.

Begley,C.G. (2013) *Nature*, **497**, 433–434.

Begley,C.G. and Ellis,L.M. (2012) *Nature*, **483**, 531–533.

Berglund,L., Björling,E. and Oksvold,P., *et al.* (2008) *Mol. Cell Proteomics*, **7**, 2019–2027.

Binz,H.K., Amstutz,P. and Plückthun,A. (2005) *Nat. Biotechnol.*, **23**, 1257–1268.

Blatt,N.B., Bill,R.M. and Glick,G.D. (1998) *Hybridoma*, **17**, 33–39.

Bordeaux,J., Welsh,A., Agarwal,S., Killiam,E., Baquero,M., Hanna,J., Anagnostou,V. and Rimm,D. (2010) *Biotechniques*, **48**, 197–209.
Bourbeillon,J., Orchard,S. and Benhar,I., *et al.* (2010) *Nat. Biotechnol.*, **28**, 650–653.
Bradbury,A.M. and Plückthun,A. (2015a) *Nature*, **518**, 27–29.
Bradbury,A.M. and Plückthun,A. (2015b) *Nature*, **520**, 295.
Bradbury,A.R.M., Sidhu,S., Duebel,S. and McCafferty,J. (2011) *Nat. Biotechnol.*, **29**, 245–254.
Collins,F.S. and Tabak,L.A. (2014) *Nature*, **505**, 612–613.
Davies,D.R., Gelinas,A.D. and Zhang,C., *et al.* (2012) *Proc. Natl. Acad. Sci. USA.*, **109**, 19971–19976.
Köhler,G. and Milstein,C. (1975) *Nature*, **256**, 495–497.
Polakiewicz,R.D. (2015) *Nature*, **518**, 483.
Ruberti,F., Cattaneo,A. and Bradbury,A. (1994) *J. Immunol. Methods*, **173**, 33–39.
Skerra,A. (2007) *Curr. Opin. Biotechnol.*, **18**, 295–304.
Slaastad,H., Wu,W., Goullart,L., Kanderova,V., Tjønnfjord,G., Stuchly,J., Kalina,T., Holm,A. and Lund-Johansen,F. (2011) *Proteomics*, **11**, 4578–4582.
Veesler,D., Dreier,B. and Blangy,S., *et al.* (2009) *J. Biol. Chem.*, **284**, 30718–30726.
Voskuil,J. (2014) *F1000Research*, **3**, 232.
Zack,D.J., Wong,A.L., Stempniak,M. and Weisbart,R.H. (1995) *Mol. Immunol.*, **32**, 1345–1353.